西湖大學
WESTLAKE UNIVERSITY

Draft Quantitative Proteomic Atlas of Human Body and Common Carcinomas

Liang Yue [1,2,3], Meng Luo [4], Wenhao Jiang [1,2,3], Ning Fan [4], Xiaolu Zhan [6], Fangfei Zhang [1,2,3], Rui Sun [1,2,3], Sainan Li [1,2,3], Tian Lu [1,2,3], Fengchao Yu [7], Guoci Teo [7], Alexey I Nesvizhskii [7,8], Ben Collins [9], Ruedi Aebersold [10,11], Fei Xu [4], Tong Liu [6], Yan Li [5], Tiannan Guo [1,2,3]

1, Key Laboratory of Structural Biology of Zhejiang Province, School of Life Sciences, Westlake University, Hangzhou, Zhejiang, China. 2, Westlake Laboratory of Life Sciences and Biomedicine, Hangzhou, Zhejiang, China 3, Institute of Basic Medical Sciences, Westlake Institute for Advanced Study, Hangzhou, Zhejiang, China 4, Department of Anatomy, College of Basic Medical Sciences, Dalian Medical University, No.9 West Section Lvshun South Road, Dalian116044, Liaoning Province, China 5, Department of Anatomy and Physiology, College of Basic Medical Sciences, Shanghai Jiao Tong University, No.280 Chongqing South Road, Shanghai, 200025, China 6, Harbin Medical University Cancer Hospital, 150 Haping Road, Harbin 150081, Heilongjiang Province, China 7, Department of Pathology, University of Michigan, Ann Arbor, Michigan, USA. 8, Department of Computational Medicine and Bioinformatics, University of Michigan, Ann Arbor, Michigan, USA. 9. Queen's University Belfast, Belfast, Northern Ireland, United Kingdom 10, Institute of Molecular Systems Biology, ETH Zurich, Zurich, Switzerland 11, Faculty of Science, University of Zurich, Zurich, Switzerland

Introduction

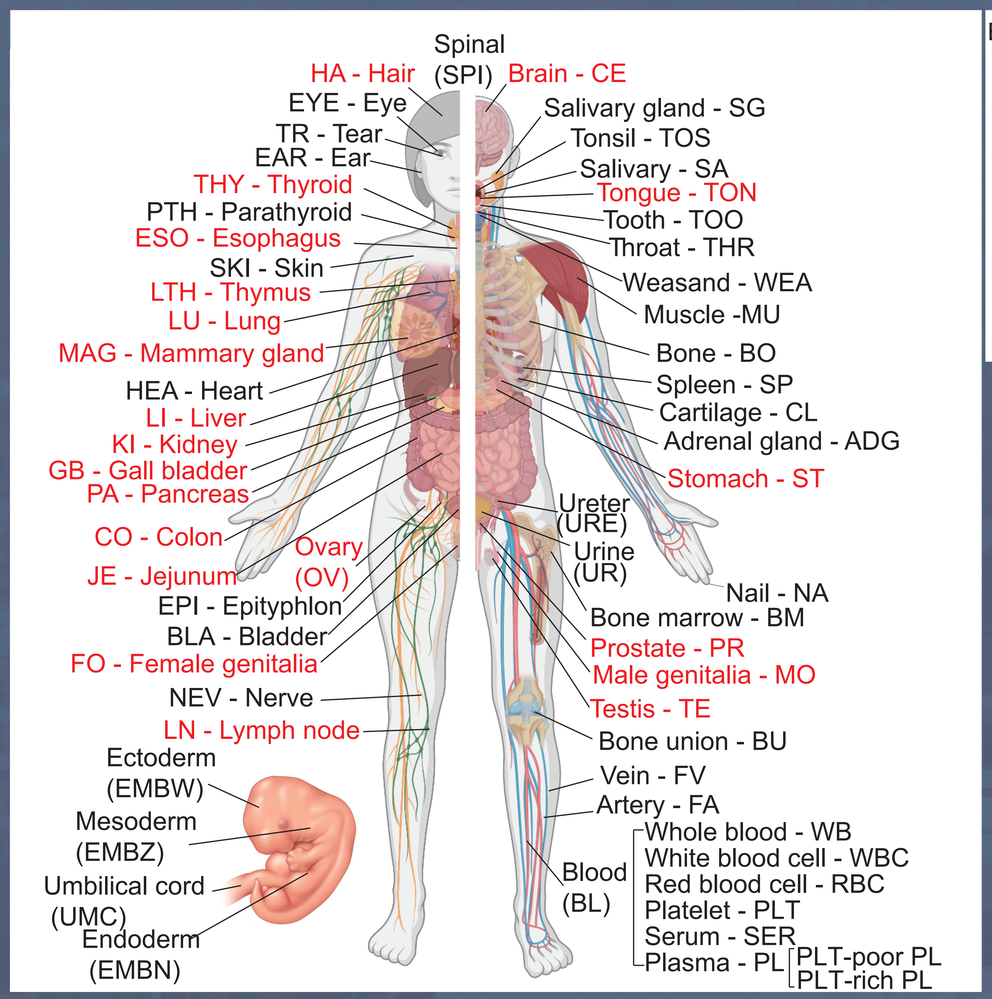


Figure 1. Overview of the sample types in TPHP

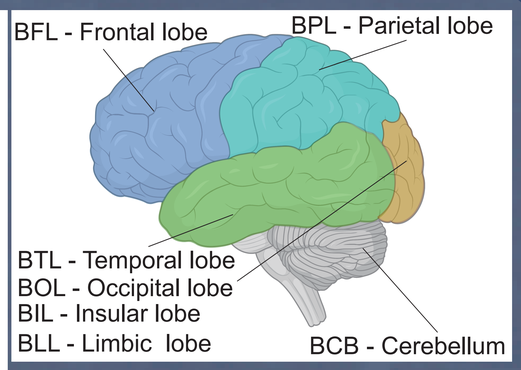
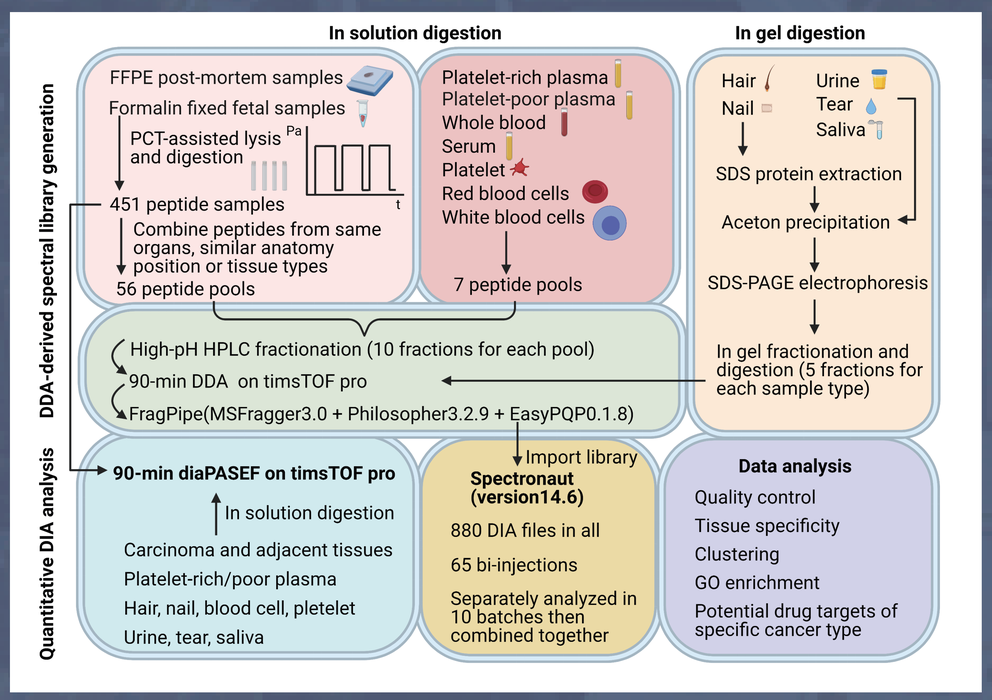


Figure 2. Brain sections

Numerous studies have investigated the proteins expressed in multiple normal and tumorous human tissues and carcinomas using qualitative or data-dependent acquisition (DDA)-based quantitative proteomics, and antibodies. However, systematic quantitative proteomic analysis with depth using a consistent method remains to be accomplished. Here, we aimed to construct a quantitative proteomic database (TPHP) from tissues all over the human body of various types and common tumors, and explore key proteins responsible for tissue specificity and malignancy.

We have collected 432 normal postmortem tissue samples from human body including 162 detailly sectioned tissues, 19 fetal tissues, and body fluids including blood, urine and saliva (Figure 1). We also collected and analyzed paired in situ carcinoma and adjacent specimens of 22 organs. Additionally, we also analyzed seven components of the blood. Seven regions from the brain were also analyzed (Figure 2).

Workflow



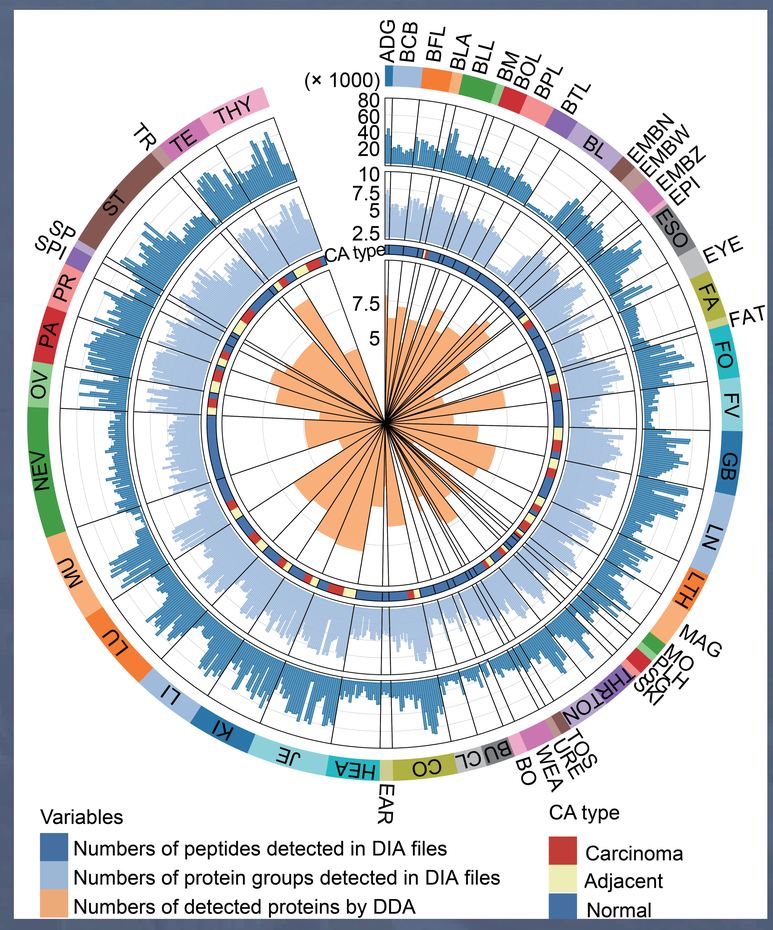
DDA derived spectral library generation

Using the pressure cycling technology (PCT)-assisted lysis and digestion [1], we processed FFPE post-mortem samples and formalin-fixed fetal samples for DDA analysis. The serum and plasma were processed by two-step overnight digestion, while the blood cells and platelet were prepared by PCT, thus producing a peptide pool for each of the seven blood components. High pH HPLC fractionation was utilized to separate the peptide samples into ten aliquots. Hair, nail, urine, tear and saliva samples were subject to acetone precipitation and in-gel digestion. Each fraction was analyzed by DDA on timsTOFpro with a 90-min gradient. Data were analyzed by FragPipe.

Quantitative DIA data analysis

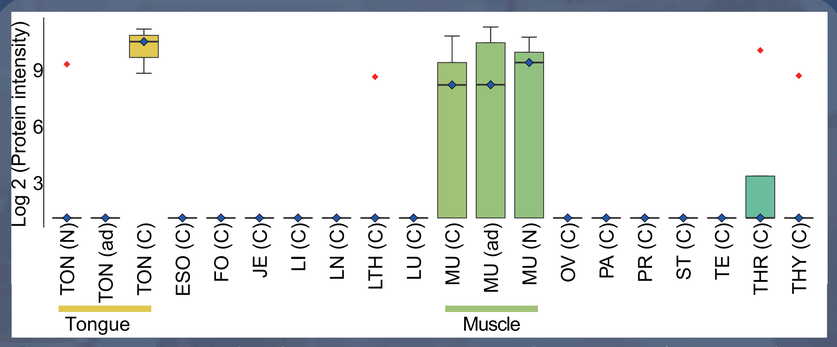
All the samples, before fractionation, for the DDA analysis were also analyzed by diaPASEF using a 90 min gradient. Moreover, over 300 paired common carcinomas and adjacent specimens were analyzed by diaPASEF too. The DIA data were analyzed by Spectronaut.

Results: Overview of TPHP proteomic data



Altogether, we analyzed 755 DDA files from 58 types of human specimens and generated 58 tissue-specific sub-libraries, including ion mobility data, using FragPipe. As shown in the radar plot in the center of the circle plot, the testis tissue library contains the highest number of proteins (8948 proteins), while the tooth library has the lowest number of proteins (1368 proteins). By combining all the 755 DDA data, the final pan-human tissue library contains 528,928 peptide precursors, 398,272 peptides and 13,668 SwissProt proteins.
With this library, 11,784 proteins were quantified in the 839 DIA files obtained in this study. The number of quantified proteins and peptides varied not only inter- but also intra- tissue types, due to the intra-organ heterogeneity in proteome composition and serious degree of pathology. In general, carcinoma and adjacent tissues led to more protein identifications than the normal tissue specimens.

Results: Potential cancer enriched markers



By comparing protein abundance among carcinoma, adjacent, normal tissues and carcinoma samples of other tissue types, we acquired a list of proteins enriched in a specific cancer type. As one example, in Figure 3, we show that CSRP3 was highly expressed in the squamous-cell carcinoma in tongue and muscle sarcoma, while less expressed or not expressed in other cancers, adjacent and normal tissue of tongue. This protein is not significantly changed between the carcinoma samples to the adjacent or normal tissues of muscle. CSRP3 is a positive regulator for the formation of muscular tissue and is also cancer enriched in head and neck cancer in Human Pathology Atlas [2]. More data analyses are ongoing.

Conclusions

1. We characterized and quantified over 10,000 proteins expressed in over 700 human samples from over 150 types of normal human tissues, body fluids, and common carcinomas.
2. We found a list of the potential cancer specific biomarkers through the comparison of quantitative proteome data of carcinoma, adjacent, relative normal tissues and carcinoma samples of other tissue type.

Acknowledgements

This work is supported by grants from the National Key R&D Program of China (2020YFE0202200), the National Natural Science Foundation of China (81972492, 21904107, 81672086, 81773022, and 82072333), Zhejiang Provincial Natural Science Foundation for Distinguished Young Scholars (LR19C050001), the Key Special Project of Ministry of Science and Technology, China (2020YFC0845700), the Fundamental Research Funds for the Central Universities (2020kfyXGYJ101), Hangzhou Agriculture and Society Advancement Program (20190101A04) and Westlake Education Foundation. We thank the support team of Bruker, FragPipe and Spectronaut for their help in acquiring and analyzing MS raw data.

References

1. Uhlen M, Zhang C, et al. A pathology atlas of the human cancer transcriptome. Science. 2017;357(6352).
2. Gao H, Zhang F, et al. Accelerated Lysis and Proteolytic Digestion of Biopsy-Level Fresh-Frozen and FFPE Tissue Samples Using Pressure Cycling Technology. J Proteome Res. 2020;19(5):1982-90.